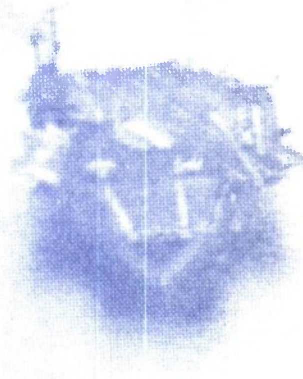
Naval Aerospace Medical Research Laboratory

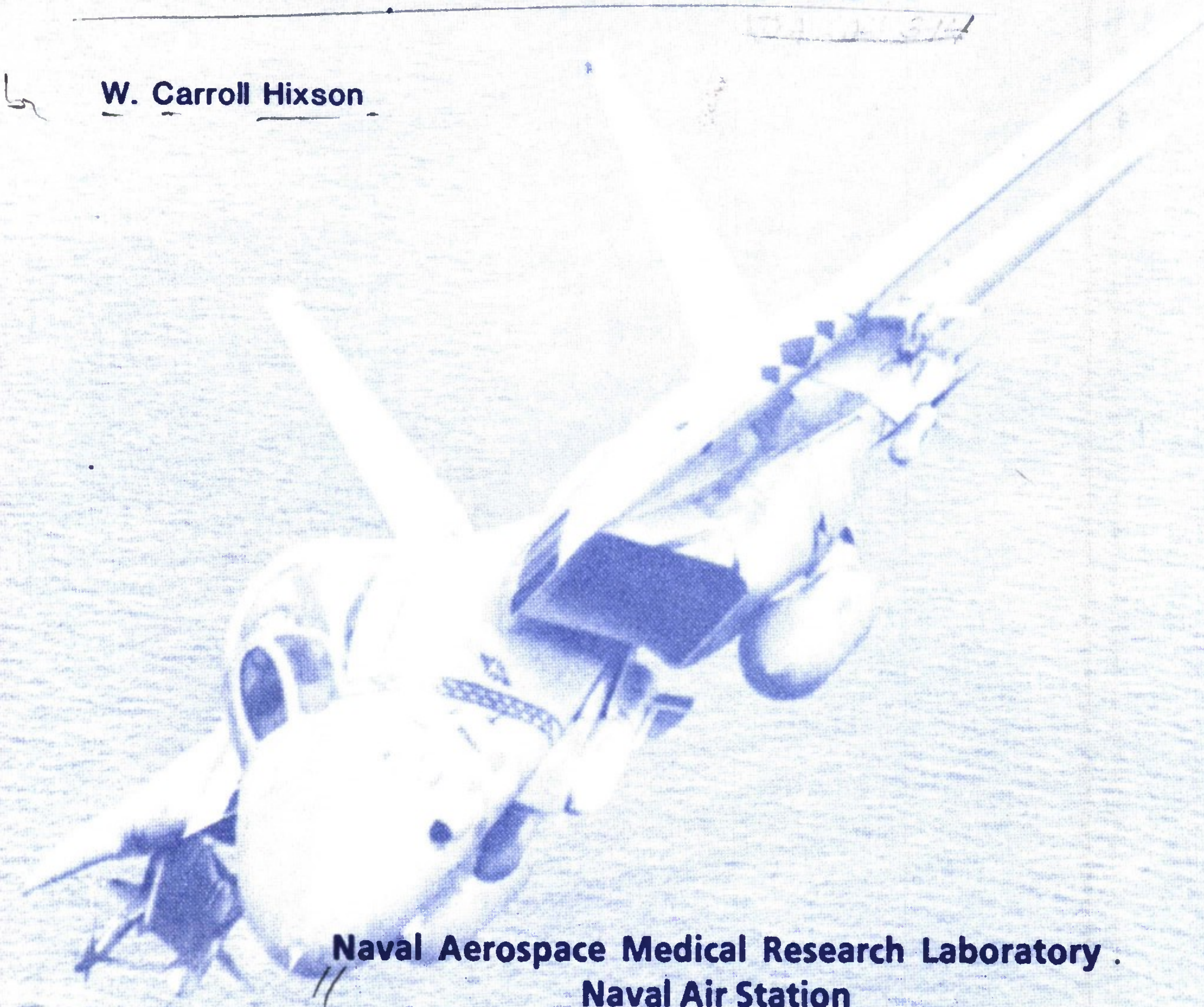
NAMRL -1345

August 1989



DEVELOPMENT OF A PERFORMANCE-BASED TEST OF GAZE CAPABILITY: A THRESHOLD APPROACH

W. Carroll Hixson



Naval Aerospace Medical Research Laboratory
Naval Air Station
Pensacola, Florida 32508-5700

ADA 214675

Reviewed and approved 16 August 1989

J. A. Brady

J. A. BRADY, CAPT, MSC USN
Commanding Officer

This research was sponsored by the Naval Medical Research and Development Command under work unit 63706N MR0096.001 DN577604.

The opinions and interpretations contained herein are those of the authors and do not necessarily represent the views, policies, or endorsement of the Department of the Navy or any other government agency.

Volunteer subjects were recruited, evaluated, and employed in accordance with the procedures specified in Department of Defense Directive 3216.2 and Secretary of the Navy Instruction 3900.39 series. These instructions are based upon voluntary informed consent and meet or exceed the provisions of prevailing national and international guidelines.

Trade names of materials and/or products of commercial or nongovernment organizations are cited as needed for precision. These citations do not constitute official endorsement or approval of the use of such commercial materials and/or products.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION
Unclassified

1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
NAMRL-1345

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION
Naval Aerospace Medical Research Laboratory

6b. OFFICE SYMBOL (If applicable)
02

7a. NAME OF MONITORING ORGANIZATION
Naval Medical Research and Development Command

6c. ADDRESS (City, State, and ZIP Code)
Naval Air Station
Pensacola, FL 32508-5700

7b. ADDRESS (City, State, and ZIP Code)
National Naval Medical Center, Bldg. 1
Bethesda, MD 20814-5044

8a. NAME OF FUNDING / SPONSORING ORGANIZATION

8b. OFFICE SYMBOL (If applicable)
Code 404

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS (City, State, and ZIP Code)
National Naval Medical Center, Bldg. 1
Bethesda, MD 20814-5044

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
63706N	MR0096	.001	DN577604

11. TITLE (Include Security Classification)
(U) Development of a Performance-based Test of Gaze Control Capability: A Threshold Approach

12. PERSONAL AUTHOR(S)
W. Carroll Hixson

13a. TYPE OF REPORT
Interim

13b. TIME COVERED
FROM Oct 86 TO Oct 88

14. DATE OF REPORT (Year, Month, Day)
89-08-16

15. PAGE COUNT
18

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Gaze, vestibular, head/eye coordination, biomedical test naval aviation

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A high degree of gaze efficiency involving the precise coordinated motion of the head and eyes is particularly important to the aviator when making large shifts in gaze during the visual scan of cockpit instruments and avionics systems. For this reason, NAMRL has concentrated on the development of a relatively low-cost, performance-based measure of head/eye coordination that can be used to identify individual differences in the naval aviation population. Preliminary research conducted by the laboratory has led to the development of a simplified test of gaze function identified as the Vestibulo-Ocular-Reflex Performance Test (VORPET). The test is based on a Bekesy-type determination of the threshold time required for an individual to recognize and identify a fixed number of digits immediately following 90° shifts in gaze. The report provides a detailed description of the new test protocol and its design concepts along with the results of experiments comparing horizontal and vertical gaze shift performance, and test-retest learning effects.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL
J. A. BRADY, CAPT MSC USN, Commanding Officer

22b. TELEPHONE (Include Area Code)
(904) 452-3286

22c. OFFICE SYMBOL
00

DD Form 1473, JUN 86 Previous editions are obsolete.

S/N 0102-LF-014-6603

SECURITY CLASSIFICATION OF THIS PAGE

DD Form 1473, JUN 86 (Reverse)

SECURITY CLASSIFICATION OF THIS PAGE

SUMMARY PAGE

THE PROBLEM

A high degree of gaze efficiency or capability, which requires the precise coördinated motion of the head and eyes, is particularly important to the aviator who must routinely make large shifts in gaze during a scan of the cockpit instruments and avionics systems. Because this capability is an important aspect of in-flight visual performance, this laboratory has initiated research to develop a relatively low-cost, performance-based measure of head/eye coordination that can be used to identify individual differences in the naval aviation population. Progress in the development of such a test has been presented in two previous reports. This report describes further progress made in test automation and simplification of the method used to gain and evaluate measures of gaze performance.

FINDINGS

Previous configurations of this test of gaze function have involved a multiple number of different exposure times with a performance score gained for each exposure time and for each direction of head movement. Although these tests met the basic objectives of this research program relative to cost and ease of operation, the extent of the information gained from the multiple performance scores probably exceeds that required for a cursory examination of gaze capability. We developed a new test, identified as the Vestibulo-Ocular-Reflex Performance Test (VORPET), that simplifies the quantified interpretation of gaze capability. The test, based on a Bekesy-type determination of the threshold time required to recognize a fixed number of digits, produces a single numerical score for each direction of head movement. The report provides a detailed description of the new test protocol and its design concepts. In addition, the results of several experiments comparing horizontal and vertical gaze shift performance and test-retest reliability are presented. The resulting data support the findings of the previous reports relative to the wide range of performance capabilities that exist within the student flight candidate population. This latter point has the potential for operational significance in that the test should distinguish pilots with exceptional gaze capabilities from those with relatively poor gaze performance.

RECOMMENDATIONS

Because the threshold concept used to develop the VORPET test provides a simplified method of distinguishing performance between individuals, it is recommended that the new test configuration be used to evaluate the gaze capabilities of different naval aviation communities.

Acknowledgments

The author wishes to acknowledge Lynn C. Percival for the original research work he conducted at NAMRL in the initial development of a performance-based test of gaze function. Thanks are also extended to Andrew N. Dennis for his contributions to the design and development of the new test configuration, and to HM2 Wendy A. Smith and HM3 Tyrone L. Green for their assistance in the collection and analysis of the test data. Similar appreciation is offered to Nell Davis for her report format contributions. Special thanks go to Fred E. Guedry, Jr., for his continued scientific input and support.

Manuscript received for review 15 October 1988.

INTRODUCTION

This laboratory has recognized the need for the development of a simple test to measure the gaze efficiency of selected members of the naval aviation community. A high degree of gaze efficiency or capability, which requires the precise coordinated motion of the head and eyes, is particularly important to the aviator who must routinely make large shifts in gaze during a scan of the cockpit instruments and avionics systems (1). It is general belief that the oculomotor and vestibulo-ocular-reflex (VOR) control systems, working in conjunction with other mechanisms (2,3), modify saccadic eye velocity automatically to maintain a high degree of head/eye coordination during large gaze shifts (4-6).

Because gaze capability is an important aspect of in-flight visual performance, we initiated research to develop a relatively low-cost, performance-based measure of head-eye coordination that can be used to identify individual differences in the naval aviation population. Progress in the development of such a test has been described in several previous reports. The original research efforts (7,8) described the theoretical aspects of the test and presented data for small and large unidirectional head motions in the horizontal plane. A follow-up paper reported progress in semiautomating the test and presented normative data for large bidirectional head movements in the vertical as well as horizontal plane.

In the initial study (7), the test procedure involved the presentation of a series of individual fixation letters on one CRT display followed by brief, time-varied presentations of digit arrays on a second CRT display located a known angular displacement from the first display. Serial letter identification was used to maintain initial gaze position. The measure of performance was the number of digits correctly identified in proper sequence following presentation of visual and auditory signals to initiate gaze shift from the fixation display to the digit array display. The second study (9) utilized light-emitting diode (LED) displays and extended the tests of gaze function to bidirectional head movements made in the vertical and horizontal head planes. For this study, performance scores were obtained for four different exposure times for each direction of head movement.

These earlier studies used test configurations involving a multiple number of fixed exposure times with a performance score gained for each exposure time and for each direction of head movement. Although these tests met the basic objectives of the research program relative to cost and ease of operation, the extent of the information gained from the multiple performance scores probably exceeds that required to identify individual differences in gaze capability.

This report provides a detailed description of a new test protocol and its design concepts. The new test, based on a Bekesy-type (10) determination of a threshold time for recognition of a fixed number of digits, produces a single numerical score for each direction of head movement. The results of several experiments involving horizontal gaze shift performance, vertical gaze shift performance, and the effects of learning on test performance are also presented.

MATERIAL AND METHODS

SUBJECTS

The subjects were volunteer, male, student naval aviators and student naval flight officers who had recently passed their flight physicals and had no known visual or vestibular deficits. A total of 72 individuals participated in the study with 30 assigned to Experiment 1, 30 to Experiment 2, and 12 to Experiment 3.

EQUIPMENT

The experimental procedures associated with the new test protocol derive in great part from the methods and techniques developed for earlier versions of the gaze function test (7-9). Similarly, the equipment used in this study is nearly identical to that used in a previous study (9). As a matter of convenience, the acronym VORPET (Vestibulo-Ocular-Reflex Performance Test) was selected to distinguish the new test protocol from its predecessor versions.

Experiments 1 and 2

The program control and data acquisition requirements for the new test configuration were provided by a desk-top microcomputer (Hewlett Packard Model 9845C). Presentation of the visual stimuli was achieved by means of four small microterminals (Burr-Brown Model TM27) located 45^{o} to the left, right, above, and below the visual dead-ahead position (see Fig. 1). Each microterminal had an eight-character LED display that utilized a hexidecimal format with character generation deriving from a seven-segment font. The characters generated with this font had the following approximate dimensions: a height of 7.62 mm (0.30 in), a width of 5.2 mm (0.2 in), and an interdigit spacing of about 5.0 mm (0.2 in) between the end of one character and the beginning of the next character. At a 1.0-m viewing distance, these dimensions resulted in viewing angles for character height, width, and interdigit spacing of approximately 26.2, 17.8, and 17.1 min, respectively. The ratio of stroke width (width of a single LED line segment) to character height was approximately 1:8.

The luminance characteristics of the LED were measured using a Pritchard spectrophotometer. As detailed in the second report (9), the luminance of a single LED line segment was measured as 17.4 cd/m^2; measurements of the luminance of the individual characters within the stimulus array characters averaged 1.36 cd/m^2.

Experiment 3

The equipment used in this experiment used the same visual display modules as described above but a different desk-top microcomputer (Hewlett Packard Series 236). The test was implemented in a newly developed Vestibular Mobile Field Laboratory constructed on a standard 8 ft by 48 ft trailer bed. Because of space limitations, the microterminals displays were located 40^{o} from the visual dead-ahead position resulting in gaze shifts of only 80^{o}.

STIMULI GENERATION

Experiments 1, 2, and 3

In these experiments, each test trial was based on the presentation of a single-character fixation stimulus on one display followed by the presentation of a four-digit numeric array on a second display located a fixed angular displacement from the first. The fixation stimulus, presented in the leftmost position of the eight-character display, consisted of the sequential display of a random number (five to nine) of dash symbols followed by a single numeric digit ranging between "2" and "9". The display of the dash symbols was marked by a 200 ms "on" period and 400 ms "off" period. The randomization process eliminated sequences that would result in repeating the number of dash symbols or the fixation digit on successive trials. The display of the single fixation digit signified to the subject that he should initiate a head movement toward the second display. This cue was reinforced by the simultaneous presentation of an audio beep tone.

In the case of the four-character numeric display, each array component utilized all digits except "0" and "1". In the randomization process, sequences that involved two adjacent digits with the same value were eliminated. In addition, components of the stimulus array that involved three sequential digits that produced a single-step run in either the up (e.g., digits 4,5,6) or down (e.g., digits 8,7,6) direction were eliminated. At the end of the preset exposure time for the array, the four stimulus digits were masked by four "0"s for a period of 500 ms.

TEST PROTOCOL

Experiment 1

This experiment measured gaze capabilities of subjects performing 90° head movements in the horizontal plane. The subject was seated in a dimly illuminated (approximately 1.5 fc) room containing the two display modules. One module was at eye level 45° to the left of the subject, and the other was 45° to the right. The 45° angular displacements were measured from the visual dead-ahead position to the leftmost digit presented on a given display.

Prior to beginning the test, the subjects were given written and verbal descriptions of the test procedures. In brief, subjects were instructed to rotate their head so as to directly face the visual display module that contained a dash symbol. The test trial was started by sequentially presenting the randomized number of dash symbols on this display. The last dash symbol was immediately followed by the display of the single numeric fixation character that occurred concurrently with an alerting audio tone. At this instant, the randomized four-digit numeric stimulus was presented on the second display for a preselected exposure time of 750 ms followed by the 500-ms visual mask. The mask was then followed by the display of a single dash character that signified to the subject that this trial had been completed and that the next test trial would start with the presentation of the fixation symbols on this second display.

Figure 1. Photograph of the VORPET test installation.

The subject task was to directly face the fixation display, watch the dash symbols, call out the single fixation digit when it appeared, and then rapidly rotate their head toward the second display and call out as many of the digits that they could identify in proper left-to-right sequence. The subjects were told that their performance would be scored as the number of digits correctly identified in proper sequence with the exception that a score of zero would be given if they did not properly identify the preceding fixation character. After each test trial, the operator manually entered the number of stimulus digits correctly identified in the computer. After entering these data, the test resumed with the alerting dash character and fixation digit appearing on the second display followed by presentation of the numeric stimulus on the first display. This procedure was continued until completion of the test.

To familiarize the subject with the presentation of the visual stimuli, each testing session was initiated by a 24-trial demonstration run that was not scored by the operator. Each trial in this run, automatically paced by the computer without operator intervention, had a fixed exposure time of 1000 ms. During this run, the test operator emphasized the need for the subject to always directly face the display where the fixation letters would be presented, to rapidly rotate his head toward the numeric stimulus display, and to call out all digits that he could correctly identify in single-digit order. Immediately following the demonstration run, a 64-trial test run was initiated with a 30-s rest period given after the first 32 trials.

Experiment 2

This experiment used the same test protocol but involved vertical head movements between two displays located 45^{o} above and 45^{o} below the visual dead-ahead position.

Experiment 3

This experiment also used the Experiment 1 test protocol with the exceptions that each subject was tested on 4 successive days, and that the two displays were located 40^{o} to the left and 40^{o} to the right of the visual dead-ahead position resulting in gaze shifts of only 80^{o}.

EXPOSURE-TIME ALGORITHM

The primary objective in developing the new test configuration was to gain a measure of the time required by an individual to correctly identify all four-digits of the numeric stimulus array. In essence, a threshold measure was desired for each individual that would identify the minimum time that the numeric stimulus array must be displayed to achieve a defined degree of performance. The approach chosen to meet this task was based on the following nomenclature. A trial was defined as a single test of gaze function involving a unidirectional head movement; a trial score was defined as the number of digits the subject correctly identified during a given trial; and a trial set was defined as eight sequential trials involving four gaze shifts made in one direction and four gaze shifts made in the opposite direction. In this test configuration, a total of 64 trials were used; 32 involving left-to-right (or up-to-down) head movements and 32

involving right-to-left (or down-to-up) head movements. Accordingly, the test was composed of eight trial sets, each involving four trials for each direction of head movement.

The test began with an initial exposure time of 750 ms for each trial. This exposure time was not adjusted until four trials (i.e., a trial set) had been completed for each direction of gaze shift. As noted in the test protocol, the operator entered the number of digits correctly identified following each trial. Since it was desired to recognize performance based on the full, rather than partial identification of all four digits in the stimulus array, a trial performance score of either one or zero was assigned following each trial: one if all four digits were identified correctly, or zero if less than four digits were identified. These scores were then used to arrive at a trial set score that could range from zero to four for the four involved trials. This trial set score was then used to increase or decrease the display exposure time for the next four trials according to the Table 1 notation.

TABLE 1. Method Used to Adjust the Time the Numeric Stimulus Array was Displayed as a Function of Subject Performance.

Trial set score	Threshold nomenclature	Display exposure time adjustment
0	Far below threshold	Increase 15 %
1	Below threshold	Increase 10 %
2	Threshold	Decrease 5 %
3	Above threshold	Decrease 10 %
4	Far above threshold	Decrease 15 %

As indicated in Table 1, threshold was arbitrarily defined as full recognition of the complete stimulus array on two of the four trials. It also could be readily argued that recognition of the full array on only one of the four trials would mark a threshold in that a single correct response indicates that the subject did indeed see and identify all four digits for the denoted exposure time. With this definition of threshold, it might be convention to leave the exposure time unchanged for the next four trials. Because learning or practice effects associated with performance tests of this form can result in a slight improvement in performance on successive trials, we decreased the exposure time a small amount whenever a threshold response occurred.

A further point is that since definitive data were not available that would define the incremental sensitivity of individuals to either variable exposure times or different degrees of gaze shift, the decision was made to linearly increase or decrease the exposure times by fixed percentage increments.

THRESHOLD NOMENCLATURE

Each 64-trial run was composed of two 32-trial segments. After the first 32 trials, the two resulting threshold times (one for each direction of head movement) were identified as the mid-thresholds. The mean of these two mid-thresholds was then calculated and stored for later analysis. The two threshold times resulted after the last 32 trials were identified as end-thresholds with the mean of the two values again calculated and stored. Although the initial exposure time for the last 32 trials began with the exposure time present at the end of the first 32 trials, the end-threshold values were determined independently of the mid-threshold values. In effect, an end-threshold could be greater than, equal to, or less than the mid-threshold. At the end of the test, the final threshold was defined as the lower of the mid- and end-threshold values for each direction of gaze shift. The simple average of the resulting two final threshold values was identified as the grand mean threshold for either horizontal or vertical gaze shifts.

RESULTS AND DISCUSSION

The results of Experiment 1, which involved 90° gaze shifts made in the horizontal head plane, are summarized in Tables 2 and 3. Table 2 compares the final threshold times obtained for the 32 left-directed head movements with the corresponding final threshold values for the 32 right-directed movements. This table presents the mean and related statistics for the final threshold times for each direction of gaze shift for the N = 30 study population; the Pearson linear correlation r coefficient between the two threshold times; and the results of a two-tailed Student t test of statistical differences between the two directions of gaze shift.

TABLE 2. Experiment 2 Horizontal Gaze Shift Data Comparing Performance Achieved with Rightward- and Leftward-Directed Gaze Shifts of 90°.

Performance statistics	Threshold times (ms) and gaze direction: Right	Left
Mean	605	600
SD	89	74
SE	16	14
Minimum	481	510
Maximum	949	863
r-correlation	0.91*	
df	28	
t-means	0.75	
df	29	

*$p < .001$

The data indicate that there were no statistically significant differences between left- and right-directed gaze performance and that a relatively strong correlation does exist between the two directions of movement. In the previous study using the same LED displays (9), statistically significant left/right directional differences were found for two of the four exposure times; performance to the right was best at the 500-ms exposure time and performance to the left was best at 1250 ms. It is probable that the pronounced differences in test protocol used in the two studies account for this apparent conflict. The Table 3 minimum and maximum data reflect the same wide range in individual performance as previously reported (7-9).

Table 3 compares the Experiment 1 mean threshold times present after the first 32 trials (mid-threshold) with those present following the last 32 trials (end-threshold). In this table, the mean data represent the simple average of the left- and right-directed threshold times at the denoted point in the test. As would be expected, there was a strong, statistically significant ($p < .001$) relationship between the two different threshold measures. The Student t statistic indicates that the end-threshold times were significantly lower than the mid-threshold times implying the presence of a learning effect over the course of the 64 trials. Although statistically significant, the numerical difference between the two scores was relatively small, with the mean end-threshold time only 38 ms less than the mid-threshold time for this study group.

TABLE 3. Experiment 1 Horizontal Gaze Shift Data Comparing Mean Threshold Times Measured After the First 32 Trials with Those Measured After the Last 32 Trials.

Performance statistics	Threshold times (ms) Mid-threshold	End-threshold
Mean	649	611
SD	81	81
SE	15	15
Minimum	558	495
Maximum	928	916
r-correlation	0.82*	
df	28	
t-means	4.24*	
df	29	

*$p < .001$

The results of Experiment 2, which involved 90° gaze shifts made in the vertical head plane, are presented in Tables 4 and 5. Table 4 compares the final threshold times measured with downward-directed gaze shifts with the corresponding thresholds measured for upward-directed gaze shifts. As

with the horizontal gaze shift data, a relatively strong correlation exists between the two directions of vertical gaze shift. Although there were no statistically significant differences in gaze performance for the two directions of horizontal gaze shift, the Table 4 data indicate that gaze performance in the upward direction was statistically better than that in the downward direction. The minimum and maximum vertical gaze shift data also reflect considerable differences in individual performance capabilities.

TABLE 4. Experiment 2 Vertical Gaze Shift Data Comparing Performance Achieved with Downward- and Upward-Directed Gaze Shifts of 90°.

Performance statistics	Threshold times (ms) and gaze direction	
	Downward	Upward
Mean	683	651
SD	83	86
SE	15	16
Minimum	539	480
Maximum	843	784
r-correlation	0.78^{**}	
df	28	
t-means	3.17^{*}	
df	29	

$^{*}p < .01$ $\quad ^{**}p < .001$

Table 5 derives from the vertical gaze shift data collected in Experiment 2 and provides a comparison between the mean mid- and end-threshold times. Again, there was a strong and statistically significant ($p < .001$) correlation between the two threshold measures. The Student *t* statistic indicates that the two mean threshold measures for the population were significantly different ($p < .001$). As with the horizontal gaze shift data, the numerical differences were relatively small, with the mean end-threshold only 40 ms less than the mid-threshold.

TABLE 5. Experiment 2 Vertical Gaze Shift Data Comparing Mean Threshold Times Measured After the First 32 Trials with Those Measured After the Last 32 Trials.

Performance statistics	Threshold times (ms)	
	Mid-threshold	End-threshold
Mean	717	676
SD	89	87
SE	16	16
Minimum	590	509
Maximum	925	835
r-correlation	0.85*	
df	28	
t-means	4.66*	
df	29	

*$p < .001$

Table 6 compares the Experiment 1 horizontal gaze shift data with the Experiment 2 vertical gaze shift data. Again, the grand mean data for the horizontal gaze shifts represent the simple average of the left- and right-directed performance scores presented in Table 2. The vertical data derives from the simple average of the downward- and upward-directed performance scores presented in Table 4. The two-tailed, independent sample, Student t statistic indicates that horizontal gaze shift performance was significantly better than vertical performance. This finding differs from that observed in the previous study (9) where no statistically significant differences were found between horizontal and vertical performance at any of four different exposure times (500, 750, 1000, and 1250 ms). However, the mean performance scores at all exposure times except 1250 ms for horizontal gaze shifts were numerically greater than the corresponding vertical values; at 1250 ms, horizontal and vertical performance was nearly identical. Assuming that horizontal gaze shifts occur more frequently than vertical gaze shifts in the normal environment, it might be projected that performance in the horizontal plane would be best as a result of greater day-to-day experience/practice. The data of the present study would lend support to this hypothesis.

TABLE 6. Comparison of Performance Scores Obtained with 90^{o} Head Movements Made in the Horizontal Plane with Those Made in the Vertical Plane.

Performance statistics	Threshold times (ms) and gaze plane	
	Horizontal	Vertical
Grand mean	602	666
SD	80	80
SE	15	15
Minimum	495	509
Maximum	606	789
t-means	-3.11*	
df	58	

*$p < .01$

The results of Experiment 3, which involved the repeat testing of 12 subjects on 4 successive days, are tabulated in Table 7. A one-way, repeated measures ANOVA of these data indicated that a significant difference (F (3, 33) = 17, $p < .001$) was present in the day-to-day treatment. In essence, a learning effect was present in that the mean threshold times decreased (performance improved) as the subjects gained experience with the test. The results of a Duncan multiple-range test comparing the four means indicated that significant improvements ($p < .05$) in performance occurred over the first 3 testing days (Table 7). However, the mean threshold time achieved on the fourth day did not differ significantly from that achieved on the third day indicating that performance stabilization was being approached. This trend can be readily visualized from Fig. 2, which is a plot of the Table 7 means and standard deviations over the 4 testing days. The

TABLE 7. Experiment 3 Data Comparing Horizontal Gaze Shift Performance (80^{o} Head Movements) for the Subject Group Tested on Four Successive Days.

Performance statistics	Threshold times (ms) and day of test			
	Day 1	Day 2	Day 3	Day 4
Grand Mean	578 a	550 b	526 c	518 c
SD	60	59	49	41
SE	17	17	14	12

Means with the same letter suffix are not significantly different according to Duncan's multiple range test ($p < .05$).

total change in threshold times over the testing period was relatively small for the subject group in that between the first and last day mean scores differed by only 60 ms.

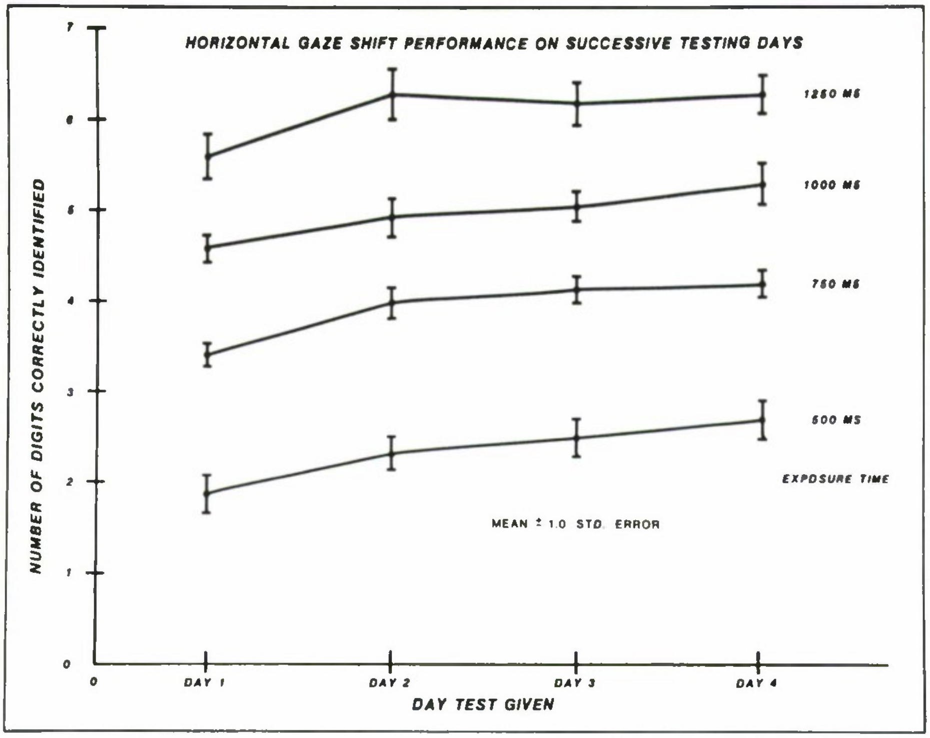


Figure 2. Plot of the Experiment 3 data showing the means and standard deviations of the threshold times measured on four-successive testing days.

The Pearson linear correlation r coefficients relating the threshold scores achieved by the subjects over the 4-day testing period are presented in matrix form in Table 8. The moderately strong correlations within the matrix indicate a good degree of test consistency.

TABLE 8. Correlation Matrix Showing the Relationships Between the VORPET Threshold Scores Obtained on Four Successive Days.

Performance statistics	Pearson correlation coefficient and day of test Day 2	Day 3	Day 4
Day 1	.91**	.86**	.86**
Day 2	---	.86**	.74*
Day 3	---	---	.73*

*p < .01 **p < .001

In the development of this VORPET protocol, a question could arise as to the absolute accuracy of the threshold measure obtained for a specific individual after only 24 practice trials and 64 test trials. It would be reasonable to expect that as a result of practice effects on such a performance test, lower threshold values would be obtained if the number of trials was increased. This is reflected in the data of Tables 3 and 4 where the mean end-thresholds measured after 64 test trials are consistently lower than the mid-thresholds arrived at after 32 trials. Although an increased number of trials or a variation in the exposure time algorithm might produce a more accurate threshold measure for an individual, little functional improvement would be gained. That is, the original project intent was to develop a short performance-based test that would both readily identify individuals with clinically significant gaze deficiencies and to distinguish individuals with exceptional gaze capabilities from those with poor capabilities. Because these data indicate a relatively wide range in gaze threshold times, increasing the duration of the test to obtain a more accurate threshold measure would probably not be a cost-effective alternative.

In actuality, it is probable that for a very cursory clinical examination, it should be possible to radically reduce the total number of trials if the primary objective is only to determine if a serious gaze deficiency exists. This could be done in various ways: 1) the algorithm could be modified to evaluate fewer trials before making the decision to change the exposure-time decision; 2) the number of digits correctly identified on a single trial could be used to adjust the exposure time for the next trial instead of the "all" or "none" criteria of the present configuration; or 3) the number of digits in the stimulus array could be reduced from 4 to some lesser number. In effect, a great deal of experimental flexibility is available to the investigator who wishes to modify the threshold determination technique developed for the present test configuration.

A last point involves the clinical role of such a test of gaze function. While the test could be used to help identify individuals with serious gaze deficiencies, the test results alone cannot identify the specific origin of the deficiency (for example, whether it arises from within the oculomotor or vestibular control mechanisms.) This identification must arise through further clinical testing of the suspected mechanisms.

RECOMMENDATIONS

As with the previous NAMRL gaze studies (7-9), the new test configuration shows considerable variation in individual performance for both horizontal and vertical gaze shifts. The range of observed performance differences lends support to the premise that VORPET has operational significance in that it should distinguish pilots with exceptional gaze capabilities from those with relatively poor gaze performance. We recommend that this version of the gaze efficiency test (VORPET) be used to acquire normative data on active squadron personnel as well as student flight candidates. It is further recommended that the VORPET test be given to a select student flight candidate population before they begin their actual in-flight training; that this population be followed as they progress through the training program; and that their performance on the VORPET be related to their flight performance during various phases of training.

REFERENCES

1. McNaughton, G.B., Aircraft Attitude Awareness Workshop, Flight Dynamics Laboratory, Wright-Patterson Air Force Base, OH, 8-10 October 1985.

2. Robinson, G.H., "Dynamics of the Eye and Head during Movement between Displays: A Qualitative and Quantitative Guide for Designers." Human Factors, Vol. 21, No. 3, pp. 343-352, June 1979.

3. Stark, L.W., "Eye-Head Coordination: Neurological Control of Active Gaze." In Space Physiology, Colloquium, Toulouse, France, 1-4 March 1983, pp. 63-75.

4. Gresty, M.A., "Coordination of Head and Eye Movements to Fixate Continuous and Intermittent Targets." Vision Research, Vol. 14, pp. 395-403, 1974.

5. Barnes, G.R., "Vestibulo-ocular Function During Coordinated Head and Eye Movements to Acquire Visual Targets." Journal of Physiology (London), Vol. 287, pp. 127-147, 1979.

6. Berthoz, A. and Melvill-Jones, G., Adaptive Mechanisms in Gaze Control, Elsevier Science Publishers, New York, NY, 1985.

7. Percival, L.C. and Guedry, F.E., Jr., Development and Evaluation of a Performance-Based Test of Saccadic and Vestibulo-Ocular Control, NAMRL-1306, Naval Aerospace Medical Research Laboratory, Pensacola, FL, January 1984.

8. Percival, L.C. and Guedry, F.E., Jr., "Performance-based Assessment of Oculomotor Efficiency." Aviation, Space, and Environmental Medicine, Vol. 59, No. 2, 1988.

9. Hixson, W.C., Guedry, F.E., Jr., and Lentz, J.M., Further Progress in Development of a Performance-based Test of Gaze Control Capability, NAMRL-1342, Naval Aerospace Medical Research Laboratory, Pensacola, FL, September, 1988.

10. von Bekesy, G., "A New Audiometer." Acta Otolaryngologica, Vol. 35, pp. 411-422, 1947.

OTHER RELATED NAMRL PUBLICATIONS*

Percival, L.C. and Guedry, F.E., Jr., Development and Evaluation of a Performance-Based Test of Saccadic and Vestibulo-Ocular Control, NAMRL-1306, Naval Aerospace Medical Research Laboratory, Pensacola, FL, January 1984. (AD A168 336)*

Percival, L.C. and Guedry, F.E., Jr., "Performance-Based Assessment of Oculomotor Efficiency." Aviation, Space, and Environment Medicine, Vol. 59, No. 2, pp. 113-120, 1988. (AD A201 391)

Hixson, W.C., Guedry, F.E., Jr., and Lentz, J.M., Further Progress in Development of a Performance-Based Test of Gaze Control Capability, NAMRL-1342, Naval Aerospace Medical Research Laboratory, Pensacola, FL, September 1988. (AD A203 250)

*These publications are available from DTIC, Cameron Station, Alexandria, VA 22314 (Phone: (C) 202/274-7633 or (AV) 284-7633). Use the AD nunber when requesting the report.